

1638

RECEIVED

APR 06 2001

TECH CENTER 1600/2900

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/549,848B

DATE: 03/27/2001

TIME: 07:25:24

Input Set : A:\17133US2.txt

Output Set: N:\CRF3\03272001\I549848B.raw

4 <110> APPLICANT: Lassner, Michael
 5 Post-Beittenmiller, Dusty
 6 Savidge, Beth
 7 Weiss, James
 9 <120> TITLE OF INVENTION: Nucleic Acid Sequences Involved in
 10 Tocopherol Synthesis
 12 <130> FILE REFERENCE: 17133/02/US
 14 <140> CURRENT APPLICATION NUMBER: 09/549,848B
 15 <141> CURRENT FILING DATE: 2000-04-14
 17 <150> PRIOR APPLICATION NUMBER: 60/129,899
 18 <151> PRIOR FILING DATE: 1999-04-15
 20 <150> PRIOR APPLICATION NUMBER: 60/146,461
 21 <151> PRIOR FILING DATE: 1999-07-30
 23 <160> NUMBER OF SEQ ID NOS: 94
 25 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 1182
 29 <212> TYPE: DNA
 30 <213> ORGANISM: Arabidopsis sp
 32 <400> SEQUENCE: 1
 33 atggagtctc tgctctctag ttcttctctt gtttccgctg ctgggtgggt ttgttggaag 60
 34 aagcagaatc taaagctcca ctctttatca gaaatccgag ttctgctgtg tgattcagat 120
 35 aaagtgtgct caaaaccgaa gtttaggaac aatcttggtt ggcctgatgg tcaaggatct 180
 36 tcattgttgt tgtatccaaa acataagtcg agatttcggg ttaatgccac tgcgggtcag 240
 37 cctgaggctt tcgactcgaa tagcaaacag aagtctttta gagactcgtt agatgcgttt 300
 38 tacaggtttt ctaggcctca tacagttatt ggcacagtgc ttagcatttt atctgtatct 360
 39 ttcttagcag tagagaaggt ttctgatata tctcctttac ttttactgg catcttgagg 420
 40 gctgttggtg cagctctcat gatgaacatt tacatagttg ggctaaatca gttgtctgat 480
 41 gttgaaatag ataaggttaa caagccctat ctccattgg catcaggaga atattctggt 540
 42 aacaccggca ttgcaatagt agcttcttcc tccatcatga gtttctggct tgggtggatt 600
 43 gttggttcat ggccattggt ctgggctctt ttgtgagtt tcatgctcgg tactgcatac 660
 44 tctatcaatt tgccactttt acggtggaaa agatttgcac tggttgcagc aatgtgtatc 720
 45 ctgctgttcc gagctattat tgttcaaatc gccttttata tacatattca gacacatgtg 780
 46 tttggaagac caatcttgtt cactaggcct ctatttttcg ccaactgcgtt tatgagcttt 840
 47 ttctctgtcg ttattgcatt gtttaaggat atacctgata tcgaagggga taagatattc 900
 48 ggaatccgat cattctctgt aactctgggt cagaaacggg tgttttggtg atgtgttaca 960
 49 ctacttcaaa tggttacgac tgttgcaatt ctagtgtgag ccacatctcc attcatatgg 1020
 50 agcaaaagta tctcggttgt gggctcatgt atactcgcaa caactttgtg ggctcgagct 1080
 51 aagtcctgtt atctgagtag caaaaccgaa ataacttcat gttatatgtt catatggaag 1140
 52 ctcttttatg cagagtactt gctgttaact tttttgaagt ga 1182
 54 <210> SEQ ID NO: 2
 55 <211> LENGTH: 393
 56 <212> TYPE: PRT
 57 <213> ORGANISM: Arabidopsis sp
 59 <400> SEQUENCE: 2
 60 Met Glu Ser Leu Leu Ser Ser Ser Ser Leu Val Ser Ala Ala Gly Gly
 61 1 5 10 15

RAW SEQUENCE LISTING

DATE: 03/27/2001

PATENT APPLICATION: US/09/549,848B

TIME: 07:25:24

Input Set : A:\17133US2.txt

Output Set: N:\CRF3\03272001\I549848B.raw

```

62 Phe Cys Trp Lys Lys Gln Asn Leu Lys Leu His Ser Leu Ser Glu Ile
63      20      25      30
64 Arg Val Leu Arg Cys Asp Ser Ser Lys Val Val Ala Lys Pro Lys Phe
65      35      40      45
66 Arg Asn Asn Leu Val Arg Pro Asp Gly Gln Gly Ser Ser Leu Leu Leu
67      50      55      60
68 Tyr Pro Lys His Lys Ser Arg Phe Arg Val Asn Ala Thr Ala Gly Gln
69 65      70      75      80
70 Pro Glu Ala Phe Asp Ser Asn Ser Lys Gln Lys Ser Phe Arg Asp Ser
71      85      90      95
72 Leu Asp Ala Phe Tyr Arg Phe Ser Arg Pro His Thr Val Ile Gly Thr
73      100     105     110
74 Val Leu Ser Ile Leu Ser Val Ser Phe Leu Ala Val Glu Lys Val Ser
75      115     120     125
76 Asp Ile Ser Pro Leu Leu Phe Thr Gly Ile Leu Glu Ala Val Val Ala
77      130     135     140
78 Ala Leu Met Met Asn Ile Tyr Ile Val Gly Leu Asn Gln Leu Ser Asp
79 145      150     155     160
80 Val Glu Ile Asp Lys Val Asn Lys Pro Tyr Leu Pro Leu Ala Ser Gly
81      165     170     175
82 Glu Tyr Ser Val Asn Thr Gly Ile Ala Ile Val Ala Ser Phe Ser Ile
83      180     185     190
84 Met Ser Phe Trp Leu Gly Trp Ile Val Gly Ser Trp Pro Leu Phe Trp
85      195     200     205
86 Ala Leu Phe Val Ser Phe Met Leu Gly Thr Ala Tyr Ser Ile Asn Leu
87      210     215     220
88 Pro Leu Leu Arg Trp Lys Arg Phe Ala Leu Val Ala Ala Met Cys Ile
89 225      230     235     240
90 Leu Ala Val Arg Ala Ile Ile Val Gln Ile Ala Phe Tyr Leu His Ile
91      245     250     255
92 Gln Thr His Val Phe Gly Arg Pro Ile Leu Phe Thr Arg Pro Leu Ile
93      260     265     270
94 Phe Ala Thr Ala Phe Met Ser Phe Phe Ser Val Val Ile Ala Leu Phe
95      275     280     285
96 Lys Asp Ile Pro Asp Ile Glu Gly Asp Lys Ile Phe Gly Ile Arg Ser
97      290     295     300
98 Phe Ser Val Thr Leu Gly Gln Lys Arg Val Phe Trp Thr Cys Val Thr
99 305      310     315     320
100 Leu Leu Gln Met Ala Tyr Ala Val Ala Ile Leu Val Gly Ala Thr Ser
101      325     330     335
102 Pro Phe Ile Trp Ser Lys Val Ile Ser Val Val Gly His Val Ile Leu
103      340     345     350
104 Ala Thr Thr Leu Trp Ala Arg Ala Lys Ser Val Asp Leu Ser Ser Lys
105      355     360     365
106 Thr Glu Ile Thr Ser Cys Tyr Met Phe Ile Trp Lys Leu Phe Tyr Ala
107      370     375     380
108 Glu Tyr Leu Leu Leu Pro Phe Leu Lys
109 385      390
111 <210> SEQ ID NO: 3

```

RAW SEQUENCE LISTING

DATE: 03/27/2001

PATENT APPLICATION: US/09/549,848B

TIME: 07:25:24

Input Set : A:\17133US2.txt

Output Set : N:\CRF3\03272001\I549848B.raw

```

112 <211> LENGTH: 1224
113 <212> TYPE: DNA
114 <213> ORGANISM: Arabidopsis sp
116 <400> SEQUENCE: 3
117 atggcgctttt ttgggctctc ccggttttca agacggttgt tgaaatcttc cgtctccgta      60
118 actccatctt ctccctctgc tcttttgcaa tcacaacata aatccttgtc caatcctgtg      120
119 actaccattt acacaaatcc ttctactaag tggtatcctt catggaatga taattaccaa      180
120 gtatggagta aaggaagaga attgcatcag gagaagtttt ttggtgttgg ttggaattac      240
121 agattaattt gtggaatgct gtcgtcttct tcggttttgg agggaaagcc gaagaaagat      300
122 gataaggaga agagtgatgg tgttggtgtt aagaaagctt ctgggataga ttgtatttta      360
123 ccagaagaag ttagagggtta tgctaagctt gctcgattgg ataaacccat tggaaacttg      420
124 ttgcttgctg ggccttgatg gtggtcgatt gcgttggtcg ctgacccctg aagccttcca      480
125 agtttttaaa atattggcttt atttggttgc ggagcattac ttcttagagg tgctggttgt      540
126 actataaatg atctgcttga tcaggacata gatacaaagg ttgatcgtac aaaactaaga      600
127 cctatcgcca gtgggtcttt gacaccattt caagggtatt gatctctcgg gctgcagttg      660
128 cttttaggct tagggattct tctccaactt aacaattaca gccgtgtttt aggggcttca      720
129 tctttgttac ttgtcttttc ctaccactt atgaagaggt ttacattttg gcctcaagcc      780
130 tttttagggt tgaccataaa ctggggagca ttgttaggat ggactgcagt taaagggaagc      840
131 atagcaccat ctattgtact cctctcttat ctctccggag tctgctggac ccttgtttat      900
132 gatactattt atgcacatca ggacaaagaa gatgatgtaa aagttggtgt taagtcaaca      960
133 gcccttagat tcggtgataa tacaaagctt tggttaactg gatttggcac agcatccata      1020
134 ggttttcttg cactttctgg attcagtgcg gatctcgggt ggcaatatta cgcactactg      1080
135 gccgctgcat caggacagtt aggatggcaa ataggacag ctgacttata atctggtgct      1140
136 gactgcagta gaaaatttgt gtcgaacaag tggtttggtg ctattatatt tagtggaagt      1200
137 gtacttgtaa gaagttttca ataa                                     1224
139 <210> SEQ ID NO: 4
140 <211> LENGTH: 407
141 <212> TYPE: PRT
142 <213> ORGANISM: Arabidopsis sp
144 <400> SEQUENCE: 4
145 Met Ala Phe Phe Gly Leu Ser Arg Val Ser Arg Arg Leu Leu Lys Ser
146 1 5 10 15
147 Ser Val Ser Val Thr Pro Ser Ser Ser Ala Leu Leu Gln Ser Gln
148 20 25 30
149 His Lys Ser Leu Ser Asn Pro Val Thr Thr His Tyr Thr Asn Pro Phe
150 35 40 45
151 Thr Lys Lys Cys Tyr Pro Ser Trp Asn Asp Asn Tyr Gln Val Trp Ser Lys
152 50 55 60
153 Gly Arg Glu Leu His Gln Glu Lys Phe Phe Gly Val Gly Trp Asn Tyr
154 65 70 75 80
155 Arg Leu Ile Cys Gly Met Ser Ser Ser Ser Val Leu Glu Gly Lys
156 85 90 95
157 Pro Lys Lys Asp Asp Lys Glu Lys Ser Asp Gly Val Val Val Lys Lys
158 100 105 110
159 Ala Ser Trp Ile Asp Leu Tyr Leu Pro Glu Glu Val Arg Gly Tyr Ala
160 115 120 125
161 Lys Leu Ala Arg Leu Asp Lys Pro Ile Gly Thr Trp Leu Leu Ala Trp
162 130 135 140
163 Pro Cys Met Trp Ser Ile Ala Leu Ala Ala Asp Pro Gly Ser Leu Pro

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/549,848B

DATE: 03/27/2001

TIME: 07:25:24

Input Set : A:\17133US2.txt

Output Set: N:\CRF3\03272001\I549848B.raw

```

164 145          150          155          160
165 Ser Phe Lys Tyr Met Ala Leu Phe Gly Cys Gly Ala Leu Leu Leu Arg
166          165          170          175
167 Gly Ala Gly Cys Thr Ile Asn Asp Leu Leu Asp Gln Asp Ile Asp Thr
168          180          185          190
169 Lys Val Asp Arg Thr Lys Leu Arg Pro Ile Ala Ser Gly Leu Leu Thr
170          195          200          205
171 Pro Phe Gln Gly Ile Gly Phe Leu Gly Leu Gln Leu Leu Leu Gly Leu
172          210          215          220
173 Gly Ile Leu Leu Gln Leu Asn Tyr Ser Arg Val Leu Gly Ala Ser
174 225          230          235          240
175 Ser Leu Leu Leu Val Phe Ser Tyr Pro Leu Met Lys Arg Phe Thr Phe
176          245          250          255
177 Trp Pro Gln Ala Phe Leu Gly Leu Thr Ile Asn Trp Gly Ala Leu Leu
178          260          265          270
179 Gly Trp Thr Ala Val Lys Gly Ser Ile Ala Pro Ser Ile Val Leu Pro
180          275          280          285
181 Leu Tyr Leu Ser Gly Val Cys Trp Thr Leu Val Tyr Asp Thr Ile Tyr
182          290          295          300
183 Ala His Gln Asp Lys Glu Asp Asp Val Lys Val Gly Val Lys Ser Thr
184 305          310          315          320
185 Ala Leu Arg Phe Gly Asp Asn Thr Lys Leu Trp Leu Thr Gly Phe Gly
186          325          330          335
187 Thr Ala Ser Ile Gly Phe Leu Ala Leu Ser Gly Phe Ser Ala Asp Leu
188          340          345          350
189 Gly Trp Gln Tyr Tyr Ala Ser Leu Ala Ala Ala Ser Gly Gln Leu Gly
190          355          360          365
191 Trp Gln Ile Gly Thr Ala Asp Leu Ser Ser Gly Ala Asp Cys Ser Arg
192          370          375          380
193 Lys Phe Val Ser Asn Lys Trp Phe Gly Ala Ile Ile Phe Ser Gly Val
194 385          390          395          400
195 Val Leu Gly Arg Ser Phe Gln
196          405
198 <210> SEQ ID NO: 5
199 <211> LENGTH: 1296
200 <212> TYPE: DNA
201 <213> ORGANISM: Arabidopsis sp
203 <400> SEQUENCE: 5
204 atgtggcgaa gatctgttgt ttctcgttta tttcaagaa tctctgtttc ttcttcgtta 60
205 ccaaacccta gactgattcc ttgggccgcg gaattatgtg ccgttaatag cttctcccag 120
206 cctccggtct cgacggaatc aactgctaag ttagggatca ctgggtgttag atctgatgcc 180
207 aatcgagttt ttgccactgc tactgccgcc gctacagcta cagctaccac cggtagagatt 240
208 tcgtctagag ttgcggcttt ggctggatta gggcatcact acgctcgttg ttattgggag 300
209 ctttctaaag ctaaacttag tatgcttggt gttgcaactt ctggaactgg gtatattctg 360
210 ggtacgggaa atgctgcaat tagcttcccg gggctttgtt acacatgtgc aggaaccatg 420
211 atgattgctg catctgctaa ttccttgaat cagatttttg agataagcaa tgattctaa 480
212 atgaaaagaa cgatgctaag gccattgcct tcaggacgta ttagtggttc acacgctggt 540
213 gcatgggcta ctattgctg tgcttctggt gcttggttgt tggccagcaa gactaatatg 600
214 ttggctgctg gacttgcac tgccaatctt gtactttatg cgtttggtta tactccgttg 660

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/549,848B

DATE: 03/27/2001

TIME: 07:25:24

Input Set : A:\17133US2.txt

Output Set : N:\CRF3\03272001\I549848B.raw

```

215 aagcaacttc accctatcaa tacatgggtt ggcgctgttg ttggtgctat cccacccttg 720
216 cttgggtggg cggcagcgtc tggtcagatt tcatacaatt cgaatgattct tccagctgct 780
217 ctttactttt ggcagatacc tcattttatg gcccttgac atctctgccg caatgattat 840
218 gcagctggag gttacaagat gttgtcactc tttgatccgt cagggaagag aatagcagca 900
219 gtggctctaa ggaactgctt ttacatgatc cctctcgggt tcatcgccca tgactggggg 960
220 ttaacctcaa gttggttttg cctcgaatca acacttctca cactagcaat cgctgcaaca 1020
221 gcattttcat tctaccgaga ccggaccatg cataaagcaa ggaaaatgtt ccatgccagt 1080
222 cttctcttcc ttcctgtttt catgtctggt cttcttctac accgtgtctc taatgataat 1140
223 cagcaacaac tcgtagaaga agccggatta acaaatcttg tatctggtga agtcaaaact 1200
224 cagaggcgaa agaaaagtgt ggtcacaact ccggtgggtt atgcctctgc tgcaaccgtt 1260
225 ctttctctcc cagctccttc cttctactct ccatga 1296
227 <210> SEQ ID NO: 6
228 <211> LENGTH: 431
229 <212> TYPE: PRT
230 <213> ORGANISM: Arabidopsis sp
232 <400> SEQUENCE: 6
233 Met Trp Arg Arg Ser Val Val Tyr Arg Phe Ser Ser Arg Ile Ser Val
234 1 5 10 15
235 Ser Ser Ser Leu Pro Asn Pro Arg Leu Ile Pro Trp Ser Arg Glu Leu
236 20 25 30
237 Cys Ala Val Asn Ser Phe Ser Gln Pro Pro Val Ser Thr Glu Ser Thr
238 35 40 45
239 Ala Lys Leu Gly Ile Thr Gly Val Arg Ser Asp Ala Asn Arg Val Phe
240 50 55 60
241 Ala Thr Ala Thr Ala Ala Ala Thr Ala Thr Ala Thr Thr Gly Glu Ile
242 65 70 75 80
243 Ser Ser Arg Val Ala Ala Leu Ala Gly Leu Gly His His Tyr Ala Arg
244 85 90 95
245 Cys Tyr Trp Glu Leu Ser Lys Ala Lys Leu Ser Met Leu Val Val Ala
246 100 105 110
247 Thr Ser Gly Thr Gly Tyr Ile Leu Gly Thr Gly Asn Ala Ala Ile Ser
248 115 120 125
249 Phe Pro Gly Leu Cys Tyr Thr Cys Ala Gly Thr Met Met Ile Ala Ala
250 130 135 140
251 Ser Ala Asn Ser Leu Asn Gln Ile Phe Glu Ile Ser Asn Asp Ser Lys
252 145 150 155 160
253 Met Lys Arg Thr Met Leu Arg Pro Leu Pro Ser Gly Arg Ile Ser Val
254 165 170 175
255 Pro His Ala Val Ala Trp Ala Thr Ile Ala Gly Ala Ser Gly Ala Cys
256 180 185 190
257 Leu Leu Ala Ser Lys Thr Asn Met Leu Ala Ala Gly Leu Ala Ser Ala
258 195 200 205
259 Asn Leu Val Leu Tyr Ala Phe Val Tyr Thr Pro Leu Lys Gln Leu His
260 210 215 220
261 Pro Ile Asn Thr Trp Val Gly Ala Val Val Gly Ala Ile Pro Pro Leu
262 225 230 235 240
263 Leu Gly Trp Ala Ala Ala Ser Gly Gln Ile Ser Tyr Asn Ser Met Ile
264 245 250 255
265 Leu Pro Ala Ala Leu Tyr Phe Trp Gln Ile Pro His Phe Met Ala Leu

```

FYI

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

DATE: 03/27/2001

PATENT APPLICATION: US/09/549,848B

TIME: 07:25:25

Input Set : A:\17133US2.txt

Output Set: N:\CRF3\03272001\I549848B.raw

L:314 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:315 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:679 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:680 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:681 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:760 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:773 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:790 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:791 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:793 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27